REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
NOV 1986

TR

4. TITLE AND SUBTITLE
SEISMIC STATION PARAMETER ESTIMATION VOL.1

5. FUNDING NUMBERS

A.P. CIERVO
G.J. HALL, JR.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
PACIFIC-SIERRA RESEARCH CORPORATION
12340 SANTA MONICA BOULEVARD
LOS ANGELES, CA. 90025

8. PERFORMING ORGANIZATION REPORT NUMBER
PSR REPORT-1552A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
ADVANCED RESEARCH PROJECTS AGENCY (ARPA)
3701 N. FAIRFAX DRIVE
ARLINGTON, VA. 22203

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
UNCLASSIFIED

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION IS UNLIMITED (A)

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

VOLUME 1 PRESENTS A PROCEDURE FOR ESTIMATING STATION SEISMICITY, NOISE, AND MAGNITUDE-BIAS PARAMETERS.

14. SUBJECT TERMS
NOISE, SEISMICITY, MAGNITUDE-BIAS PARAMETERS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT

18. SECURITY CLASSIFICATION OF THIS PAGE

19. SECURITY CLASSIFICATION OF ABSTRACT

20. LIMITATION OF ABSTRACT

PSR Report 1552A

SEISMIC STATION PARAMETER ESTIMATION

Vol. I

A. P. Ciervo
G. J. Hall, Jr.

Revised
November 1986

Sponsored by
U.S. Arms Control and Disarmament Agency
Washington, D.C. 20451

PACIFIC-SIERRA RESEARCH CORPORATION
12340 Santa Monica Boulevard • Los Angeles, California 90025 • (213) 820-2200
A subsidiary of Eaton Corporation

PREFACE

This report is part of a continuing research effort sponsored by the Defense Advanced Research Projects Agency (DARPA) and the U.S. Arms Control and Disarmament Agency (ACDA) to resolve technical issues concerning verification of nuclear test ban treaties. Volume I of this report presents a procedure for estimating station seismicity, noise, and magnitude-bias parameters. The noise parameters are required inputs for the Seismic Network Assessment Program for Detection (SNAP/D). Volume II applies the procedure developed in Vol. I to the AEDS classified seismic network.

An earlier identically titled version of this report (PSR Report 1552, Vols. I and II, August 1985) derived station parameter estimates without accounting for the effect of the network detection criteria on station histograms. This version accounts for this effect and extends the results of the earlier report to station-amplitude histograms as well as station m_b histograms.

CONTENTS

I. INTRODUCTION

This report develops a model that can be used in conjunction with single-station histogram data to estimate station seismicity and performance parameters. The histogram data generally consists of a plot of the number of earthquakes (from a restricted epicentral region) versus station m_b, although a model for the treatment of station amplitude histograms for worldwide seismic data is also considered. The estimated station parameters consist of the mean noise level, μ, the standard deviation of log noise, σ_n, and the station magnitude bias, e. In order to estimate e, a similar histogram of network detection performance corrected for maximum-likelihood (ML) m_b magnitudes must also be available.

The noise parameters μ and σ_n for each station are required inputs for the Seismic Network Assessment Program for Detection (SNAP/D) [Ciervo, et al., 1983]. Previously, their values were either estimated from measurements made on seismograms, or inferred from station detection thresholds. The empirical method does not generally ensure accurate replication of station performance in SNAP/D runs and the threshold-inferred estimates are generally reliable only for magnitudes near the station threshold. The procedure presented here is based on past station performance throughout the magnitude range experienced by the station and is thus believed to be an improvement on prior noise parameter estimation procedures.

A similar treatment for a somewhat different problem has been presented by Kelly and Lacoss [1969] where ML estimates were derived for network performance parameters. However, for mathematical convenience, a single-station Gaussian detection model was used to represent the network detection process. Furthermore, the biasing effect of non-ML corrected network m_b estimates [see Ringdal 1976] was not

understood at that time. A similar effect on single-station seismicity estimates is accounted for in the procedure derived here. In addition, single-station noise estimates are corrected for the usual four-station P-wave network detection criteria.

The procedures developed here are illustrated using histogram data for the station HYB (Hyderabad, India) observing events from the Kamchatka/Kurile region of the USSR. The estimation procedures are also applied to the analysis of classified station m_b histograms as detailed in Vol. II of this report. However, station M_s histogram data was either too sparse or irregular to obtain reliable estimates, hence Vol. II presents M_s noise parameters from a previously published report [Hutchenson, 1983].

II. NOTATION

The following notation will be adopted for the discussion below:

m = operational m_b

$\hat{m}$ = m_b observed at a single station

a = log amplitude (log A/T)

$\hat{a}$ = log amplitude observed at a single station

$N(m)$ = density for the expected number of earthquakes occuring with magnitude m

α = intercept of base e seismicity

β = slope of base e seismicity

μ = station mean noise amplitude

e = station magnitude bias

r = SNR required for station observation

$b(\Delta)$ = b-factor (i.e., $m = \log (A/T) - b(\Delta)$)

μ' = $-b(\Delta) + \log \mu + \log r$

Φ = unit normal probability distribution

σ_s = standard deviation (s.d.) of $\hat{m}$ given m

σ_n = s.d. of single-station log noise

y_k = number of events in $\underline{k}$th magnitude interval of station histogram

i = station index

j = epicentral region index

III. SEISMICITY MODEL

It is generally accepted that logarithmic seismicity from a given region is linear with respect to seismic magnitude [Richter, 1958]. Thus, if $N(m)\Delta$ is the average number of seismic events per year occurring in the operational magnitude* interval $(m - \Delta/2, m + \Delta/2)$, then define α and β such that

$$N(m) = e^{\alpha+\beta m} \quad . \tag{1}$$

As in Kelly and Lacoss [1969], the actual number of events in an operational magnitude interval of width Δ is assumed to be Poisson distributed with mean $N(m)\Delta$. Although, to the best of our knowledge, no formal justification has been offered for this assumption, it is reasonable since over a fixed time interval (say, one year), the occurrence of primary earthquakes parameterized on magnitude appear to satisfy the axioms of a nonhomogeneous Poisson process [Parzen, 1962].

Define the kth operational magnitude interval as $(m_k - \Delta/2, m_k + \Delta/2)$, where $\Delta = m_{k+1} - m_k$, $k = 1, 2, \ldots$, and X_k as the random number of earthquakes with operational magnitude within the kth interval. Then the Poisson assumption implies that

$$\mathcal{P}\{X_k = x\} = e^{-N(m_k)\Delta} \frac{[N(m_k)\Delta]^x}{x!} \quad x = 0, 1, 2, \ldots \tag{2}$$

where $N(m_k)$ is given by Eq. (1).

At this point no claim has been made about the distribution of Y_k, the number of earthquakes detected by a single station with observed magnitude in the interval $(\hat{m}_k - \Delta/2, \hat{m}_k + \Delta/2)$. However, using Eq. (2) and the results below, Appendix A proves that Y_k is also Poisson.

---

*For a discussion of true, operational, and observed magnitudes see von Seggern and Blandford [1976]. Unless otherwise noted, all magnitudes are m_b values with the subscript suppressed.

IV. SINGLE-STATION DETECTION

The amplitude of a seismic signal arriving at a station may be considered to result from a series of random multiplicative (attenuation) effects on the seismic source amplitude. The central limit theorem would then imply that the log of the station amplitude, and hence observed magnitude $\hat{m}$, is a Gaussian random variable given the operational magnitude m. Thus,

$$\mathcal{P}\{\hat{m}|m\} = \frac{1}{\sigma_s\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\hat{m}-m}{\sigma_s}\right)^2} \tag{3}$$

where σ_s is the log signal s.d.

The log of the station noise amplitude at any time is also Gaussian since the noise is generally composed of signals from myriad minor seismic disturbances. Suppose a station detects a seismic signal with amplitude s when s/n > r where n is the noise amplitude and r is the signal-to-noise ratio (SNR) required for detection. The probability of detection would then be

$$P_D = \mathcal{P}\{s/n > r\}$$

$$= \mathcal{P}\{\log s - (\log n + \log r) > 0\} \quad . \qquad (4)^*$$

From the discussion above, log s and log n are Gaussian with expectation and variance given by (in SNAP/D notation [Ciervo, et al., 1983])

$$E(\log s) = \log(A/T)$$
$$V(\log s) = \sigma_s^2$$
$$E(\log n) = \log \mu$$

*The following notation is used: $\log_{10} = \log$ and $\log_e = \ell n$.

and

$$V(\log n) = \sigma_n^2 \quad .$$

where A is the mean signal amplitude in nm at the dominant wave period T. Seismologists prefer to use the quantity A/T because of its relationship to the energy in the wave train [Richter, 1958].

Thus, Eq. (4) becomes

$$P_D = \Phi\left[\frac{\log(A/T) - (\log \mu + \log r)}{\sqrt{\sigma_s^2 + \sigma_n^2}}\right] \tag{5}$$

which is essentially Eq. (6) in the SNAP/D User's Manual. The relationship between magnitude m and amplitude A is

$$m = \log(A/T) - b(\Delta) \tag{6}$$

where b is the correction factor for epicentral distance Δ. Defining

$$\mu' = -b(\Delta) + \log \mu + \log r \tag{7}$$

Eqs. (6) and (7) allow Eq. (5) to be rewritten as

$$P_D = \Phi\left(\frac{m - \mu'}{\sqrt{\sigma_s^2 + \sigma_n^2}}\right) \quad . \tag{8}$$

Note that Eq. (8) is actually the probability of single-station detection conditioned on operational magnitude m. It is also useful to consider the detection probability conditioned on the observed magnitude $\hat{m}$. In this case, the only uncertainty is the noise amplitude so that [Von Seggern and Blandford, 1976],

$$\mathcal{P}\{\mathcal{D} \mid \hat{m}\} = \Phi\left(\frac{\hat{m} - \mu'}{\sigma_n}\right) \tag{9}$$

where $\mathcal{D}$ denotes detection.

V. STATION HISTOGRAM MODEL

Incremental histogram data is generally a plot of y_k = the number of events detected with observed magnitude in the interval $(\hat{m}_k - \Delta/2, \hat{m}_k + \Delta/2)$ where $\Delta = \hat{m}_{k+1} - \hat{m}_k$, $k = 1, 2, \ldots$. The histogram data y_k is a realization of the random variable Y_k discussed on p. 4. The expectation of Y_k, which is needed for station parameter estimation, is derived below from the seismicity and single-station detection models above.

Recalling that N(m) is the average density of earthquakes at operational magnitude m, the average density of earthquakes arriving at a station with observed magnitude $\hat{m}$ is

$$\int_0^\infty \mathcal{P}\{\hat{m}|m\}\ N(m)\ dm \quad .$$

Thus, the average density of earthquakes detected by a single station is, using Eqs. (1), (3), and (9), given by

$$\begin{aligned}\hat{N}(\hat{m}) &= \mathcal{P}\{\mathcal{D}|\hat{m}\} \int_0^\infty \mathcal{P}\{\hat{m}|m\}\ N(m)\ dm \\ &= \Phi\left(\frac{\hat{m}-\mu'}{\sigma_n}\right) \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma_s}\ e^{-\frac{1}{2}\left(\frac{\hat{m}-m}{\sigma_s}\right)^2} e^{\alpha+\beta m}\ dm \\ &\cong e^{\alpha'+\beta\hat{m}}\ \Phi\left(\frac{\hat{m}-\mu'}{\sigma_n}\right) \qquad (10)\end{aligned}$$

where $\alpha' = \alpha + 0.5\ \beta^2\sigma_s^2$ and the approximation is due to the negligible effect of using 0 instead of $-\infty$ for the lower limit of the integral. Thus, as noted by von Seggern and Blandford [1976], the apparent

effect of using station histogram data to estimate seismicity is to introduce an upward bias of 0.5 $\beta^2\sigma_S^2$ into the intercept parameter α.

The expectation of Y_k is given by

$$E(Y_k) = \hat{N}(\hat{m}_k)\,\Delta$$

$$= 10^{A'+B\hat{m}}\,\Phi\left(\frac{\hat{m}-\mu'}{\sigma_n}\right) \qquad (11)$$

where, from Eq. (10), $B = (\ell n\ 10)^{-1}\beta$ and $A' = (\ell n\ 10)^{-1}\ (\ell n\Delta + \alpha')$. The unbiased histogram intercept is then $A = (\ell n\ 10)^{-1}\ (\ell n\ \Delta + \alpha)$ so that the upward bias is $A' - A = 0.5\ (\ell n\ 10)^{-1}\beta^2\sigma_S^2 = 0.5(\ell n\ 10)B^2\sigma_S^2$. Using the assumption that operational seismicity is Poisson, Appendix A proves that the Y_k are Poisson random variables so that Eq. (11) is also the variance of Y_k.

The above discussion assumes that detection at an individual station is independent of the performance of the remaining stations in the network. In fact, the teleseismic detection of an earthquake generally requires at least a four-station detection of the P-wave. The effect of ignoring network influence is to overestimate station noise levels. To properly account for the effect of the network detection criteria on the station i histogram, the probability of the remaining stations in the network (i.e., the "reduced network") providing at least a three-station P-wave detection must be considered. With this approach, Eq. (10) becomes

$$\hat{N}(\hat{m}) = \mathcal{P}\{\mathcal{D}_S\,|\hat{m}\} \int_0^{\infty} \mathcal{P}\{\mathcal{D}_N\,|m\}\,\mathcal{P}\{\hat{m}|m\}\,N(m)\,dm\ . \qquad (12)$$

where $\mathcal{D}_S$ and $\mathcal{D}_N$ represent detection by station i and the reduced network, respectively. Since a closed form expression for $\mathcal{P}\{\mathcal{D}_N|m\}$ does not exist, SNAP/D is run over a range of m's for each reduced network. The integral in Eq. 12 is then computed numerically.

Note that each SNAP/D run in the numerical evaluation of Eq. 12 requires the noise inputs μ and σ for each station in the reduced network. This requires an iterative process which begins with zeroth order μ and σ estimates* (from the assumption of independent station histograms) to determine $\mathcal{P}\{\mathcal{D}_N|m\}$ for each reduced network, which in turn are used to calculate first order μ and σ estimates, and so on. The detailed results of such an iterative process are shown in Vol. II of this report for the AEDS network. After five iterations, the estimated values for μ', σ, and A converged to fifth digit accuracy. The resulting μ values converged to at least three digit accuracy.

* The parameter estimation procedure is described in Section VII.

VI. AMPLITUDE HISTOGRAMS

The expressions derived in the previous section can be generalized to station i log amplitude histograms provided the event's epicentral region, j, is known, along with the seismicity of the region. Denoting the quantity, log A/T, as 'a' for brevity, Eq. (6) yields

$$a = m + b(\Delta) \quad .$$

The attenuation factor, b, depends on the angular distance Δ_{ij} between station i and epicenter j, which is indicated in the following development by b_{ij}. Noting that the observed $\hat{m}_i$ is related to the operational m_i by a normally distributed deviation δ_{ij}, with mean e_{ij} and standard deviation σ_{si}, the above expression becomes

$$\hat{a}_i = m_i + b_{ij} + \delta_{ij} \quad .$$

Eq. (12) can be now be rewritten as

$$\hat{N}(\hat{a}_k) = \sum_j \left[\int_m N_j(m)\, \mathcal{P}\{\mathcal{D}_N | m \text{ at } j\}\ \mathcal{P}\{\hat{a}_k | m \text{ at } j\}\ \mathcal{P}\{\mathcal{D}_S | \hat{a}_k\} \right] \tag{13}$$

where

$$N_j(m) = e^{\alpha_j + \beta_j m}$$

$$\mathcal{P}\{\hat{a}_k | m \text{ at } j\} = \frac{1}{\sqrt{2\pi}\,\sigma_{si}}\, e^{-1/2\left(\frac{\hat{a}_k - b_{ij} - e_{ij} - m}{\sigma_{si}}\right)^2}$$

and

$$\mathcal{P}\{\mathcal{D}_S | a_k\} = \Phi\left(\frac{\hat{a}_k - b_{ij} - \mu_i'}{\sigma_i}\right) \quad .$$

The necessary inputs for such a general expression would be α_j, β_j, e_{ij}, and σ_{si}. The integral still must be evaluated numerically with $\mathcal{P}\{\mathcal{D}_N | m \text{ at } j\}$ computed by SNAP/D.

There are three special cases for the amplitude histogram problem:

1. $\beta_j = \bar{\beta}$ for all j: with this assumption, $\bar{\beta}$ can be estimated as in the earlier discussion by treating data as independent and averaging the fitted β_j's. The SNAP/D iterations and the numerical integral would still be required.
2. Independent amplitude histograms are equivalent to

$$\mathcal{P}\{\mathcal{D}_N | m \text{ at } j\} \equiv 1.0$$

and would remove the need for the SNAP/D iterations.

3. Independent amplitude histograms with $\beta_j = \bar{\beta}$ would cause Eq. 13 to become

$$\hat{N}_i(\hat{a}_k) = \sum_j \left[\int_m N_j(m)\, \mathcal{P}\{\hat{m} \text{ at } i | m \text{ at } j\}\, \mathcal{P}\{\mathcal{D}_S | \hat{m} \text{ at } i\} \right] dm$$

$$= \Phi\left(\frac{\hat{a}_k - \mu_i''}{\sigma_i}\right) \sum_j \int_m e^{\alpha_j + \beta_j m}\; e^{-1/2\left(\frac{\hat{m} - m - e_{ij}}{\sigma_{si}}\right)^2}$$

$$\times \frac{dm}{\sqrt{2\pi}\,\sigma_{si}} \qquad (14)$$

Where

$$\mu_i'' = \log \mu_i + \log r_i$$

Defining

$$\alpha_j' = \alpha_j + \frac{\beta^2\, \sigma_{si}^2}{2} - \beta\, e_{ij}$$

and

$$\alpha' = \ell n \left\{ \sum_j e^{\left[\alpha_j - \beta\left(b_{ij} + e_{ij}\right) + \frac{\beta^2 \sigma_{si}^2}{2}\right]} \right\} = \ell n \left[\sum_j e^{\alpha_j' - \beta b_{ij}} \right],$$

simplifies Eq. (14) to

$$\hat{N}_i(\hat{a}_k) = e^{\alpha' + \beta \hat{a}_k} \quad \Phi\left(\frac{\hat{a}_k - \mu_i''}{\sigma_i}\right) \tag{15}$$

so that the functional form of the expected number of events in a bin is similar to the independent m_b case Eq. (10). In Eq. (10), α', μ_i'', and σ_i would generally be regarded as free parameters to be estimated, while β may be regarded as fixed or free.

VII. PARAMETER ESTIMATION

For ease of computation, a minimum chi-square (MCS) estimation procedure was chosen. However, Appendix B proves the asymptotic (large Σy_k) equivalence of MCS and ML estimates for this problem. For each station the MCS estimates are those that minimize the usual chi-square sum, i.e.,

$$\left(\min_{\mu', \sigma_n, A', B}\right) \sum_k \frac{(o_k - e_k)^2}{e_k} \tag{16}$$

where, in our case, the kth observation $o_k = y_k$, and the expected observation $e_k = N(\hat{m}_k)$, where N is given by Eq. (10) or (12) depending on whether independent or dependent station data is used . In practice, the MCS sum is computed over all k for which $e_k \geq 1$, where the parameter values at each stage of the iterative minimization process are used to compute e_k. The technical properties of the MCS estimates (and their asymptotic equivalence to maximum likelihood (ML) estimates) are detailed in Appendix B where the variance of the estimates are also derived.

Estimates for each station are calculated in three steps: (1) the minimization indicated in Eq. (16) is performed (under the assumption of independent station histograms) with all four parameters free; (2) the weighted average

$$\bar{B} = \frac{1}{n} \sum_i n_i B_i$$

is calculated where B_i is the station i slope estimate from step (1), n_i is the number of events in the station i histogram and $n = \Sigma n_i$; and (3) the minimization is repeated with three free parameters and $B = \bar{B}$

fixed for all stations. In the case of network dependent histograms, the iterative minimization procedure presented in Sec. V must be used. The rationale for setting $B = \overline{B}$ to obtain the final μ', σ_n, and A' estimates is that $\overline{B}$ is generally a better estimate of the operational slope than an individual B_i.

SNAP/D requires mean station noise levels in amplitude units (0-P in nm) and standard deviations in log amplitude units. Thus, if the minimization procedure described above provides μ'_i and σ_{ni} estimates for station i, then the SNAP/D station i noise inputs are, from Eq. (7),

$$\mu_i = 10^{\,\mu'_i + b(\Delta_i) - \log r}$$

and σ_{ni} unchanged.

The complications arising from network influences on m_b histograms, as discussed in Sec. V, imply that, in this case, an unbiased seismicity estimate may not be available from the MCS fit. However, the results of applying the procedure to AEDS station data (presented in Vol. II of this report) indicate a negligible change in A' parameter estimates. Thus, the expression for the unbiased station i intercept parameter, $A_i = A'_i - 0.5(\ell n\, 10)B^2\sigma^2$, as discussed in Sec. V for the case of independent m_b data, is also used as a reasonable approximation for the case of dependent data. Since SNAP/D calculations are conditioned on the occurrence of a seismic event, seismicity is not a SNAP/D input for unassociated data, but estimation of the intercept parameter A_i permits estimation of station magnitude bias as discussed below.

If network estimates of operational seismicity parameters A_{NET} and B_{NET} are available, then the MCS procedure would consist only of step (3) with B_i set equal to B_{NET}. In this case, the estimate of station magnitude bias would be

$$e_i = \frac{A_{NET} - A_i}{B_{NET}} \quad ,$$

this expression assumes that the station i was in operation continuously. If the periods for which station i is inoperable are known, an obvious modification of A_{NET} could be made so that the e_i calculation would still be correct. The usual caution associated with using network histogram data to estimate seismicity must be observed: only MLE corrected network magnitude data can be used to obtain unbiased estimates of A_{NET} and B_{NET} due to the "bulge" phenomena associated with histograms based on network average magnitudes [Zavadil, et al., 1983]. Although SNAP/D can in principle accommodate corrections for station magnitude bias through use of the correction factor e_{ijk} (in SNAP/D notation: i = station index, j = epicenter index, and k = wave index), in practice the required data acquisition and analysis would be formidable even when restricted to P-wave observation for seismically active areas in the Soviet Union.

VIII. EXAMPLE

The calibration procedure described above was applied to 1976-1980 histogram data compiled by F. Ringdal [Rivers, 1984] for Kamchatka events observed by the station in Hyderabad, India. The application assumes independent station data for simplicity of discussion. This is clearly an approximation to the true case of network association. Figure 1 plots the data, the MCS fit (Eq. (6)), the unbiased station seismicity ($10^{A_i+\bar{B}m}$), and, since no ML network magnitude data was available, hypothetical unbiased network seismicity ($10^{A_{NET}+B_{NET}m}$). Assuming $B_{NET} = \bar{B}$, $\Delta = 0.1$, and $\sigma_s = 0.35$, the station magnitude bias $0.5(\ln 10)B^2\sigma_s^2$ is also indicated. Note that the apparent station seismicity bias refers to the "vertical" difference between apparent asymptotic (large magnitude) station seismicity ($10^{A'_i+\bar{B}m}$) and the unbiased station seismicity ($10^{A_i+\bar{B}m}$). On the other hand, the station magnitude bias refers to the "horizontal" difference between the unbiased station seismicity and the unbiased network seismicity ($10^{A_{NET}+B_{NET}m}$).

The MCS estimates for HYB are

$$\mu' = 4.83$$
$$\sigma_n = 0.20$$
$$A' = 6.15$$

with seismicity slope fixed at B = −0.89. Thus, SNAP/D noise parameters would be

$$\mu = 10^{\mu'} + b(\Delta) - \log r$$

$$= 7.48$$

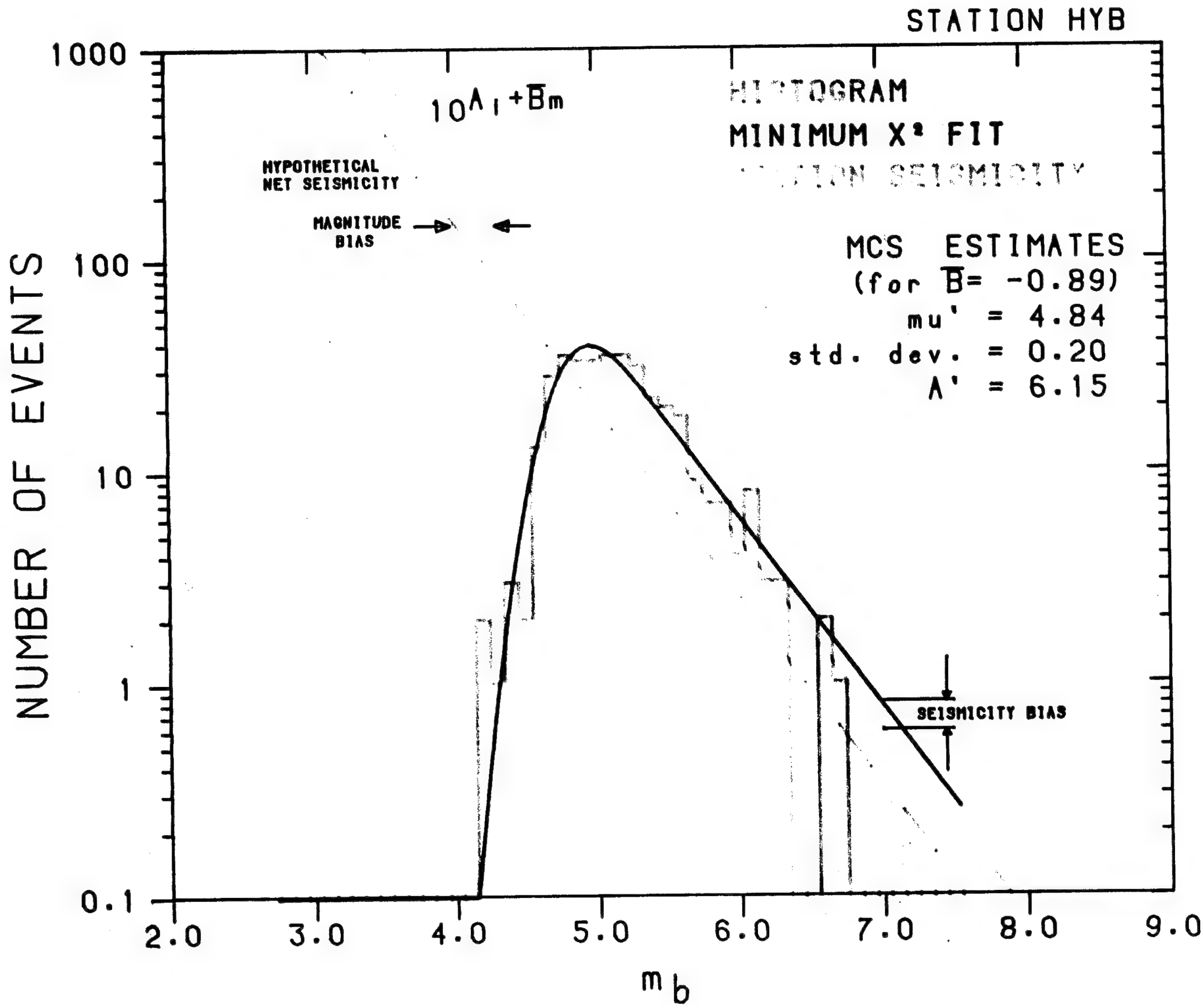

Figure 1. Illustration of calibration fit and parameters.

and $\sigma_n = 0.20$ where $b(\Delta) = 3.78$ is the Veith and Clawson [1972] 0 km entry for $\Delta = 68.82°$ and $r = 1.5$. The unbiased seismicity intercept is

$$A = A' + 0.5(\ell n\ 10)B^2\sigma_s^2$$

$$= 6.04 \quad .$$

Appendix A

THE DISTRIBUTION OF OBSERVED MAGNITUDES

This appendix will show that if $X(m)$, the random number of earthquakes with operational magnitude in the interval $[0,m)$, is a Poisson process, then $Y_D(\hat{m})$, the number of earthquakes detected by a single station with observed magnitude in $[0,\hat{m})$, is also Poisson. Hence, the histogram data Y_k discussed on p. 4 of the text is a Poisson random variable.

Let $X(m)$ be a Poisson process with density function

$$\lambda(m) = e^{\alpha-\delta m} \quad ,$$

where $\delta > 0$, $m \in [0,\infty)$, and $\delta = -\beta$, in the notation of Sec. III. $X(m)$ denotes the total number of earthquakes with operational magnitude in the interval $[0,m)$, and $\overline{X}(m)$ is the total number of earthquakes with magnitude bigger than m (i.e., magnitude in the interval (m,∞))*. Both $X(m)$ and $\overline{X}(m)$ are assumed to be Poisson processes, with respective mean value functions

$$E[X(m)] \triangleq \Lambda(m)$$

$$= \int_0^m \lambda(s)\, ds$$

$$= \frac{e^{\alpha}}{\delta}\left(1 - e^{-\delta m}\right)$$

*$\overline{X}(m)$ is the Poisson process treated by Kelly and Lacoss [1969].

$$E[\bar{X}(m)] \triangleq \bar{\Lambda}(m)$$

$$= \int_m^\infty \lambda(s)\, ds$$

$$= \frac{e^{\alpha-\delta m}}{\delta}$$

In an interval $\left(m - \frac{\Delta}{2}, m + \frac{\Delta}{2}\right)$, the expected number of earthquakes is

$$\Lambda\left(m + \frac{\Delta}{2}\right) - \Lambda\left(m - \frac{\Delta}{2}\right) = \int_{m-\frac{\Delta}{2}}^{m+\frac{\Delta}{2}} e^{\alpha-\delta s}\, ds$$

$$\cong e^{\alpha-\delta m}\, \Delta$$

Now to develop the distribution of $Y(\hat{m})$, the number of earthquakes with <u>observed</u> magnitude in $[0,\hat{m})$, let η_i be the operational magnitude of the i^{th} smallest earthquake recorded by the station, where η_i may be smaller or bigger than $\hat{m}$. Let $\hat{\eta}_i$ denote the corresponding observed magnitude, so that

$$\hat{\eta}_i = \eta_i + \varepsilon_i \ ,$$

where the measurement error ε_i is Gaussian with mean 0 and standard deviation σ_s.

Define the 0-1 valued function

$$w(\hat{m},\eta,\varepsilon) = \begin{cases} 1 & \text{if} \quad \eta + \varepsilon \le \hat{m} \\ 0 & \text{if} \quad \eta + \varepsilon > \hat{m} \end{cases}$$

Then the i^{th} earthquake is included in the $Y(\hat{m})$ count if $\hat{\eta}_i \le \hat{m}$, i.e., if $w(\hat{m}, \eta_i, \varepsilon_i) = 1$, and otherwise not. Thus, $Y(\hat{m})$ may be expressed as

$$Y(\hat{m}) = \sum_{i} w(\hat{m}, \eta_i, \varepsilon_i)$$

Clearly, $Y(\hat{m})$ is a nonnegative integer-valued random variable.

The assertion that $Y(\hat{m})$ is again a Poisson process can be proved in several different ways, but we shall present a proof that is straightforward and intuitively appealing.

First, we demonstrate a fairly well known lemma [Papoulis, 1984] relating the limiting distribution of a sum of independent Bernoulli (0-1 valued) random variables to a Poisson variate.

Let $x_1, x_2, \ldots, x_n, x_{n+1}, \ldots,$ be a sequence of independent random variables such that x_i is 1 with probability p_i and 0 with probability $q_i = 1 - p_i$. Further, let

$$\lambda = \lim_{n \to \infty} \sum_{i=1}^{n} p_i < \infty$$

and assume that $\max\limits_{1 \le i \le n} p_i \to 0$ as $n \to \infty$.

Then we assert that $Z_n = \sum_{i=1}^{n} x_i$ converges (as $n \to \infty$) to a Poisson random variable Z with mean λ.

To show this, we exhibit the characteristic function of each x_i and of Z, and show consequently that the characteristic function of Z_n converges to that of Z, thus establishing the lemma.

The characteristic function of the Bernoulli variable x_i is

$$\phi_i(u) = E\left(e^{jux_i}\right)$$

$$= p_i\ e^{ju} + q_i$$

(where $j = \sqrt{-1}$) and the characteristic function of the Poisson variate Z is

$$\phi_Z(u) = E\left(e^{juZ}\right)$$

$$= \sum_{k=0}^{\infty} e^{juk} \frac{e^{-\lambda}\lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^{ju}\right)^k}{k!}$$

$$= e^{\lambda(e^{ju}-1)}$$

Now note that if $p_i << 1$, then

$$e^{p_i(e^{ju}-1)} \cong 1 + p_i(e^{ju}-1)$$

$$= p_i e^{ju} + q_i$$

$$= \phi_i(u) \qquad \text{(A.1)}$$

and so

$$\ell n\, \phi_{Z_n}(u) = \sum_{i=1}^{n} \ell n\, \phi_i(u) \qquad \text{(by independence of } x_i\text{)}$$

$$= \sum_{i=1}^{n} \ell n\left(p_i e^{ju} + 1 - p_i\right)$$

$$= \sum_{i=1}^{n} \left[p_i\left(e^{ju} - 1\right) + \theta_i p_i\right]$$

$$= \sum_{i=1}^{n} p_i\left(e^{ju} - 1\right) + \sum_{i=1}^{n} \theta_i\, p_i$$

$$\xrightarrow{n \to \infty} \lambda(e^{ju} - 1) \qquad \text{(A.2)}$$

where $\theta_i \to 0$ as $p_i \to 0$. Hence $\phi_{Z_n}(u) \longrightarrow \phi_Z(u)$, establishing the lemma.

To show that if $X(m)$ is Poisson, then so is $Y(\hat{m})$, begin by partitioning the magnitude axis $[0,\infty)$ into consecutive intervals $I_i = (\alpha_i, \alpha_{i+1})$ of fixed length (or mesh) $\Delta\alpha = \alpha_{i+1} - \alpha_i$, as in Fig. A.1. Let ΔX_i denote the number of earthquakes with operational magnitude in the interval I_i and let $\Delta Y(\hat{m}, \alpha_i)$ denote the corresponding contribution to the sum $Y(\hat{m})$ due to the error in measuring the operational magnitudes of the earthquakes in interval I_i. Thus, we may write $Y(\hat{m})$ as

$$Y(\hat{m}) = \sum_i \Delta Y(\hat{m}, \alpha_i)$$

Figure A.1

If the mesh $\Delta\alpha$ of the partition is small, then, within probabilities of order $\Delta\alpha$, the random variable ΔX_i takes the value 0 or 1, and

$$P(\Delta X_i = 1) \cong \lambda(\alpha_i)\ \Delta\alpha$$

where, as before, $\lambda(m) = e^{\alpha-\delta m}$ is the intensity function of the Poisson process $X(m)$.

If $\Delta X_i = 0$, then necessarily $\Delta Y(\hat{m}, \alpha_i) = 0$. But if $\Delta X_i = 1$, then

$$\Delta Y(\hat{m}, \alpha_i) = 1 \longleftrightarrow \hat{\eta}_{\ell_i} = \eta_{\ell_i} + \varepsilon_{\ell_i} \leq \hat{m}$$

$$\text{or} \quad \varepsilon_{\ell_i} \leq \hat{m} - \eta_{\ell_i}$$

where the operational magnitude η_{ℓ_i} lies in interval I_i.

Thus, approximately, given that $\Delta X_i = 1$,

$$P\left[\Delta Y(\hat{m},\alpha_i) = 1 \mid \Delta X_i = 1\right] \cong \Phi\left(\frac{\hat{m} - \alpha_i}{\sigma_s}\right)$$

since the measurement error is Gaussian with mean 0 and standard deviation σ_s. Hence the unconditional probability is

$$P\left[\Delta Y(\hat{m}, \alpha_i) = 1\right] \cong \lambda(\alpha_i)\ \Delta\alpha\ \Phi\left(\frac{\hat{m} - \alpha_i}{\sigma_s}\right)$$

Now, the random variables $\Delta Y(\hat{m}, \alpha_i)$ take the values 0 or 1 and are independent, because the ΔX_i are independent (Poisson variates in nonoverlapping intervals) and the measurement errors ε_{ℓ_i} are independent. Therefore, by the lemma, as the mesh $\Delta\alpha \to 0$, the sum $Y(\hat{m}) = \sum_i \Delta Y(\hat{m}, \alpha_i)$ tends to a Poisson variate with mean

$$\lim_{\Delta\alpha\to 0} \sum_i \lambda(\alpha_i)\ \Phi\left(\frac{\hat{m} - \alpha_i}{\sigma_s}\right) \Delta\alpha = \int_0^\infty \lambda(m)\ \Phi\left(\frac{\hat{m} - m}{\sigma_s}\right) dm \qquad (A.3)$$

The above argument also illustrates that for two nonoverlapping intervals $(0, \hat{s})$ and $(\hat{m}, \hat{m} + t)$, where $\hat{s} < \hat{m}$, the variates $Y(\hat{s})$ and $Y(\hat{m} + t) - Y(\hat{m})$ are independent, and, of course, Poisson. Hence $Y(\hat{m})$ is also a Poisson process, as was to be shown.

From Eq. (A.3), we may write the mean value function for $Y(\hat{m})$ as

$$M(\hat{m}) = \int_0^\infty \lambda(m)\Phi\left(\frac{\hat{m} - m}{\sigma_s}\right) dm$$

and the corresponding intensity function is

$$\mu(\hat{m}) = \frac{d}{d\hat{m}} M(\hat{m}) = \int_0^\infty e^{\alpha+\beta m} \frac{1}{\sigma_s\sqrt{2\pi}} e^{-\frac{(\hat{m}-m)^2}{2\sigma_s^2}} dm$$

$$= \int_0^\infty N(m) \mathscr{P}\{\hat{m}|m\} dm , \tag{A.4}$$

as in Section V, pg. 7.

Finally, we consider the effect of the probability of detection

$$\mathscr{P}\{\mathscr{D}|\hat{m}\} = \Phi\left(\frac{\hat{m} - \mu'}{\sigma_n}\right)$$

on the process $Y(\hat{m})$. As before, partition the magnitude axis into disjoint intervals $\hat{I}_i = (\hat{\alpha}_i, \hat{\alpha}_{i+1})$ with mesh $\Delta\hat{\alpha} = \hat{\alpha}_{i+1} - \hat{\alpha}_i$, and let $\Delta\hat{\alpha}$ be so small that within probabilities of order $\Delta\hat{\alpha}$ the number ΔY_i of occurrences of earthquakes with <u>measured</u> magnitude in $\hat{I}_i$ is 0 or 1. If $\Delta Y_i = 0$, define $\Delta Y_D(\hat{m}, \hat{\alpha}_i) = 0$.

If $\Delta Y_i = 1$, define

$$\Delta Y_D(\hat{m}, \hat{\alpha}_i) = 1 \longleftrightarrow \text{the event in } \hat{I}_i \text{ is \underline{detected}.}$$

Thus, given $\Delta Y_i = 1$, the conditional probability that $\Delta Y_D(\hat{m}, \hat{a}_i)$ is 1 is

$$P\left[\Delta Y_D\ (\hat{m}, \hat{a}_i) = 1 \mid \Delta Y_i = 1\right] = \mathcal{P}\{\mathcal{D}|\hat{m}\} \tag{A.5}$$

and as developed earlier

$$P(\Delta Y_i = 1) \cong \mu(\hat{a}_i)\ \Delta\hat{a}\ . \tag{A.6}$$

The total number of earthquakes <u>detected</u> in $[0, \hat{m})$ is

$$Y_D(\hat{m}) = \sum_i \Delta Y_D(\hat{m}, \hat{a}_i)\ .$$

And from Eqs. (A.5) and (A.6),

$$P\left[\Delta Y_D(\hat{m}, \hat{a}_i) = 1\right] \cong \mu(\hat{a}_i)\mathcal{P}\{\mathcal{D}|\hat{a}_i\}\ \Delta\hat{a}_i$$

Thus, as the mesh $\Delta\hat{a}\to 0$, the total number of earthquakes with magnitude in $[0, \hat{m})$ that are detected is a Poisson process with mean

$$M_D(\hat{m}) = \int_0^{\hat{m}} \mu(x)\ \mathcal{P}\{\mathcal{D}|x\}\ dx \tag{A.7}$$

In an interval $\left(\hat{m} - \frac{\Delta}{2}, \hat{m} + \frac{\Delta}{2}\right)$ the number of earthquakes that are detected is given by

$$M_D\left(\hat{m} + \frac{\Delta}{2}\right) - M_D\left(\hat{m} - \frac{\Delta}{2}\right) = \int_{\hat{m}-\frac{\Delta}{2}}^{\hat{m}+\frac{\Delta}{2}} \mu(x)\ \mathcal{P}\{\mathcal{D}|x\}\ dx$$

$$\approx \mathcal{P}\{\mathcal{D}|\hat{m}\} \int_{\hat{m}-\frac{\Delta}{2}}^{\hat{m}+\frac{\Delta}{2}} \mu(x)\, dx$$

Finally, by differentiating Eq. (A.7), we obtain the density (see Eq. (A.4) for $\mu(\hat{m})$) for the detected process,

$$\begin{aligned}\mu_D(\hat{m}) &= \mu(\hat{m})\, \mathcal{P}\{\mathcal{D}|\hat{m}\} \\ &= \mathcal{P}\{\mathcal{D}|\hat{m}\} \int_0^\infty N(m)\, P\{\hat{m}|m\}\, dm \\ &= \hat{N}(\hat{m}) \ ,\end{aligned}$$

as in Eq. (10) of Sec. V of the text.

Appendix B

MINIMUM CHI-SQUARE AND MAXIMUM LIKELIHOOD IN FITTING A POISSON PROCESS MODEL

In this appendix, we assume that count observations can fall into any one of K bins or magnitude intervals, where each bin has the same width Δ. Let

$$
\begin{aligned}
y_k &= \text{number of observed counts in } k^{th} \text{ bin} \\
e_k &= \text{expected number of counts in } k^{th} \text{ bin} \\
&= \lambda_k(\theta) \\
&= e^{\alpha+\beta m_k} \Phi\left(\frac{m_k - \mu}{\sigma_n}\right) \\
&= e^{\alpha}\, h(m_k,\, \theta') \\
&\equiv \gamma h(m_k,\, \theta)
\end{aligned}
$$

for each k, $k = 1, \ldots, K$, where m_k is the midpoint of the kth bin, $\theta = (\beta, \mu, \sigma_n)$, $\theta' = (\alpha, \theta)$, and the definition of $h(.)$ is clear from the above. Note $\gamma \equiv e^{\alpha}$.

Assuming that $y_1, \ldots, y_K$ are independent Poisson random variables, where y_k has mean λ_k, then $N = \Sigma y_k$ is also Poisson with mean $\Sigma\lambda_k$. Moreover, the joint density function (or likelihood function) of $y_1, \ldots, y_K$ is

$$L(\theta') = f(y_1, \ldots, y_K) = \frac{\prod_k e^{-\lambda_k} \lambda_k^{y_k}}{\prod_k y_k!} \tag{B.1}$$

The density function for the sum N is

$$P(N = n) = \frac{e^{-\sum_k \lambda_k} \left(\sum_k \lambda k\right)^n}{n!}, \qquad \text{for } n = 0, 1, \ldots .$$

The log of the likelihood function from Eq. (B.1) is then

$$\log L = -\sum_k \lambda_k + \sum_k y_k \ln \lambda_k - \sum_k \ln y_k!$$

$$= -\sum_k \gamma h(m_k, \theta) + \sum_k y_k \left[\ln \gamma + \ln h(m_k, \theta)\right] - \sum_k \ln y_k! \tag{B.2}$$

MAXIMUM LIKELIHOOD ESTIMATION (MLE)

The maximum likelihood estimate (MLE) $\hat{\theta}'$ for θ' is obtained by choosing $\hat{\theta}'$ to maximize $L(\theta')$, or equivalently $\log L(\theta')$. If $L(\theta')$ is differentiable with respect to θ' (as in our model), the MLE $\hat{\theta}'$ is obtained by solving the equation

$$\frac{\partial}{\partial \theta'} \log L(\theta') = 0 . \tag{B.3}$$

By differentiating Eq. (B.2) with respect to γ, we obtain

$$0 = \frac{\partial \log L}{\partial \gamma} = -\sum_k h(m_k, \theta) + \frac{1}{\gamma} \sum_k y_k ,$$

or

$$\hat{\gamma} = \frac{\sum_k y_k}{\sum_k h(m_k, \theta)} = \frac{N}{\sum_k h(m_k, \theta)} \tag{B.4}$$

Substituting this expression for γ in the log likelihood (B.2), we obtain (aside from a constant depending on N, but not depending on θ) the log likelihood for the conditional distribution of $(y_1, \ldots, y_K)$ given $N = \sum_k y_k$, which is multinomial with parameters N, $\pi_1(\theta), \ldots, \pi_K(\theta)$, where for $1 \leq k \leq K$,

$$\pi_k(\theta) = \frac{h(m_k,\ \theta)}{\sum_j h(m_j,\ \theta)} = \frac{\lambda_k(\theta')}{\sum_j \lambda_j(\theta')}$$

This version of the log likelihood may then be maximized (by solving for $\partial \log L/\partial\theta = 0$) to obtain the MLE estimates for $\theta = (\beta,\ \mu,\ \sigma_n)$.

MINIMUM CHI-SQUARE ESTIMATION (MCSE)

Although the asymptotic properties of MCS are well known for the multinomial case [Cox and Hinkley, 1977 and Rao, 1957], they do not seem to be not as well known for the Poisson case. Since we were not able to find in the literature an exact reference for the Poisson case, we present the development here.

For the Poisson model, the MCS estimate is achieved by minimizing

$$S(\theta') = \sum_k \frac{(y_k - e_k)^2}{e_k}$$

$$= \sum_k \frac{(y_k - \lambda_k(\theta'))^2}{\lambda_k(\theta')} \tag{B.5}$$

where e_k is the expected number of counts in bin k, which in this model is $\lambda_k(\theta')$.

The MCSE is obtained by solving the equation

$$\frac{\partial}{\partial\theta'} S(\theta') = 0 \quad .$$

We will show that under suitable regularity conditions, the solution of

$$\frac{\partial}{\partial\gamma} S(\gamma) = 0$$

yields an estimate $\tilde{\gamma}$ which is asymptotically equivalent to the ML

estimate $\hat{\Upsilon}$, and that substitution of the resulting estimate into Eq. (B.5) yields (approximately) the MCSE for the multinomial distribution, which is known to be asymptotically equivalent to the MLE.

Now the MCS criteria is

$$S(\theta') = \sum_k \frac{\left[y_k - \Upsilon h(m_k,\ \theta)\right]^2}{\Upsilon h(m_k,\ \theta)} \tag{B.6}$$

Letting $h_k = h(m_k,\ \theta)$ and differentiating with respect to Υ, we obtain

$$0 = \frac{\partial S}{\partial \Upsilon} = \sum_k \left[\frac{-2(y_k - \Upsilon h_k)\ h_k^2\ \Upsilon - (y_k - \Upsilon h_k)^2 h_k}{\Upsilon^2\ h_k^2}\right]$$

$$= \sum_k \left(\frac{\Upsilon^2 h_k^2 - y_k^2}{\Upsilon^2 h_k}\right),$$

hence

$$\tilde{\Upsilon} = \left(\frac{\sum_k \frac{y_k^2}{h_k}}{\sum_j h_j}\right)^{1/2}$$

is the MCS estimate of Υ, for fixed θ.

The ML estimate $\hat{\Upsilon}$ is unbiased, since

$$E\left(\hat{\Upsilon}\right) = E\left(\frac{N}{\sum_k h_k}\right) = \frac{\Upsilon \sum_k h_k}{\sum_k h_k} = \Upsilon .$$

Its variance (since N is Poisson) is

$$\text{Var } \hat{\gamma} = \text{Var}\left(\frac{N}{\sum_k h_k}\right)$$

$$= \frac{1}{\left(\sum_k h_k\right)^2} \gamma \sum_k h_k$$

$$= \frac{\gamma}{\sum_k h_k}$$

The variance becomes small if $\sum_k h_k$ is large, which we shall assume.

To evaluate the mean of the MSCE γ, let

$$Z = \frac{\sum_k \frac{y_k^2}{h_k}}{\sum_j h_j} .$$

Since y_k is Poisson with parameter $\gamma\ h_k$,

$$E\ (y_k^2) = \gamma^2\ h_k^2 + \gamma\ h_k \quad .$$

Thus

$$E(Z) = \frac{\sum_k \left(\frac{\gamma^2 h_k^2 + \gamma h_k}{h_k}\right)}{\sum_j h_j}$$

$$= \gamma^2 + \gamma \frac{K}{\sum_k h_k}$$

$$= \gamma^2 \left(1 + \frac{K}{\gamma \sum_k h_k}\right)$$

To first order, then, since $\tilde{\gamma} = Z^{1/2}$,

$$E\,\tilde{\Upsilon} \cong \Upsilon \left(1 + \frac{K}{\Upsilon \sum_k h_k}\right)^{1/2}$$

$$\cong \Upsilon \left(1 + \frac{K}{2\,\Upsilon \sum_k h_k}\right)$$

where the last approximation follows from the Taylor expansion of $\sqrt{1+X}$.

If $\frac{K}{\sum_k h_k}$ goes to 0 (or $K^{-1} \sum_k h_k$ grows large) the bias term goes to zero.

It may also be shown, using the first four moments of the Poisson distribution, that to first order,

$$\mathrm{Var}\,\tilde{\Upsilon} \cong \frac{\Upsilon}{\sum_k h_k} \quad ,$$

just as for the MLE $\hat{\Upsilon}$.

Thus, if K and $\sum_k h_k$ grow large in such a way that $\Upsilon/\sum_k h_k$ and $K/\sum_k h_k$ decreases toward zero, $\tilde{\Upsilon}$ and $\hat{\Upsilon}$ will be asymptotically equivalent.

Writing

$$\tilde{\Upsilon} = \hat{\Upsilon} + \varepsilon(\Upsilon)$$

$$= \frac{N}{\sum_k h_k} + \varepsilon(\Upsilon)$$

where the error $\varepsilon(\Upsilon)$ depends on Υ and on $\Sigma\, h_k$, we replace Υ by $\tilde{\Upsilon} = N/\sum_k h_k + \varepsilon(\Upsilon)$ in Eq. (B.6), and obtain

$$S(\theta) = \sum_k \frac{\left\{y_k - \left(N\frac{h_k}{\sum_j h_j} + \varepsilon(\Upsilon)\, h_k\right)\right\}^2}{\frac{N\, h_k}{\sum_j h_j} + \varepsilon(\Upsilon)\, h_k}$$

But if $|\varepsilon(\Upsilon)\, h_k|$ is small for each k, this yields (approximately) the MCS criterion for the multinomial distribution of $(y_1 \ldots, y_K)$ given N. From Ferguson [1958], the resulting MCS estimate for $\theta = (\beta, \mu, \sigma_n)$ is asymptotically equivalent (as N grows large) to the ML estimate (i.e., $\tilde{\theta}$ is Best Asymptotically Normal, or BAN).

The asymptotic equivalence of MCS and ML estimates allow us to use the usual MSE approach [Cox and Hinkley, 1974] to evaluate the asymptotic variance for the MCS estimates. Thus, we evaluated

$$- \frac{\partial^2}{\partial \theta'^2} \log L(\theta') \Bigg|_{\theta' = \tilde{\theta}'}$$

where $\tilde{\theta}'$ is the MCSE, and inverted this matrix to achieve the estimated variance-covariance matrix for θ'. Elements of this matrix are given below (with a sign reversal):

$$\frac{\partial^2(\log L(\theta'))}{\partial \alpha^2} = \sum_i \left\{-e^{\alpha+\beta m_i}\ \Phi\left(\frac{m_i - \mu}{\sigma_n}\right)\right\}$$

$$\frac{\partial^2(\log L(\theta'))}{\partial \alpha\ \partial \beta} = \sum_i \left\{-m_i\ e^{\alpha+\beta m_i}\ \Phi\left(\frac{m_i - \mu}{\sigma_n}\right)\right\}$$

$$\frac{\partial^2(\log L(\theta'))}{\partial \alpha\ \partial \mu} = \sum_i \left\{\frac{1}{\sigma_n}\ e^{\alpha+\beta m_i}\ \phi\left(\frac{m_i - \mu}{\sigma_n}\right)\right\}$$

$$\frac{\partial^2(\log L(\theta'))}{\partial\alpha\,\partial\sigma_n} = \sum_i \left\{ \frac{m_i-\mu}{\sigma_n^2}\, e^{\alpha+\beta m_i}\ \phi\left(\frac{m_i-\mu}{\sigma_n}\right)\right\}$$

$$\frac{\partial^2(\log L(\theta'))}{\partial\beta^2} = \sum_i \left\{ -m_i^2\, e^{\alpha+\beta m_i}\ \Phi\left(\frac{m_i-\mu}{\sigma_n}\right)\right\}$$

$$\frac{\partial^2(\log L(\theta'))}{\partial\beta\,\partial\mu} = \sum_i \left\{ \frac{m_i}{\sigma_n}\, e^{\alpha+\beta m_i}\ \phi\left(\frac{m_i-\mu}{\sigma_n}\right)\right\}$$

$$\frac{\partial^2(\log L(\theta'))}{\partial\beta\,\partial\sigma_n} = \sum_i \left\{ \frac{m_i(m_i-\mu)}{\sigma_n^2}\, e^{\alpha+\beta m_i}\ \phi\left(\frac{m_i-\mu}{\sigma_n}\right)\right\}$$

$$\frac{\partial^2(\log L(\theta'))}{\partial\mu^2} = \sum_i \left\{ \frac{m_i-\mu}{\sigma_n^3}\, e^{\alpha+\beta m_i}\ \phi\left(\frac{m_i-\mu}{\sigma_n}\right) - \frac{y_i}{\sigma_n^2}\, \frac{\phi\left(\frac{m_i-\mu}{\sigma_n}\right)}{\Phi\left(\frac{m_i-\mu}{\sigma_n}\right)} \left[\frac{m_i-\mu}{\sigma_n} + \frac{\phi\left(\frac{m_i-\mu}{\sigma_n}\right)}{\Phi\left(\frac{m_i-\mu}{\sigma_n}\right)}\right]\right\}$$

$$\frac{\partial^2(\log L(\theta'))}{\partial\mu\,\partial\sigma_n} = \sum_i \left\{ -\frac{1}{\sigma_n^2}\, e^{\alpha+\beta m_i}\ \phi\left(\frac{m_i-\mu}{\sigma_n}\right)\left[1-\left(\frac{m_i-\mu}{\sigma_n}\right)^2\right] + \frac{y_i}{\sigma_n^2}\, \frac{\phi\left(\frac{m_i-\mu}{\sigma_n}\right)}{\Phi\left(\frac{m_i-\mu}{\sigma_n}\right)} \left[1-\left(\frac{m_i-\mu}{\sigma_n}\right)^2 - \frac{\phi\left(\frac{m_i-\mu}{\sigma_n}\right)}{\Phi\left(\frac{m_i-\mu}{\sigma_n}\right)}\left(\frac{m_i-\mu}{\sigma_n}\right)\right]\right\}$$

$$\frac{\partial^2(\log L(\theta'))}{\partial\,\sigma_n^2} = \sum_i \left\{ \frac{m_i-\mu}{\sigma_n^3}\, e^{\alpha+\beta m_i}\ \phi\left(\frac{m_i-\mu}{\sigma_n}\right)\left[-2+\left(\frac{m_i-\mu}{\sigma_n}\right)^2\right] + \frac{y_i(m_i-\mu)}{\sigma_n^3}\, \frac{\phi\left(\frac{m_i-\mu}{\sigma_n}\right)}{\left(\frac{m_i-\mu}{\sigma_n}\right)} \left[2 - \frac{m_i-\mu}{\sigma_n}\left(\frac{m_i-\mu}{\sigma_n} + \frac{\phi\left(\frac{m_i-\mu}{\sigma_n}\right)}{\left(\frac{m_i-\mu}{\sigma_n}\right)}\right)\right]\right\}$$

In summary, we have shown that

1. The MLE $\hat{\Upsilon}$ is unbiased for Υ, with variance inversely proportional to $\Sigma\, h_k$; thus Υ converges to Υ in probability if $\Sigma\, h_k$ grows large.
2. The MCSE $\tilde{\Upsilon}$ is biased, with bias

$$b(\theta) = \frac{K}{2 \sum_k h_k}$$

As $K^{-1} \sum_k h_k$ grows large, the bias decreases toward zero. The variance of $\tilde{\Upsilon}$ is, to first order $\Upsilon/\Sigma\, h_k$, just as for the MLE.
3. We may write $\tilde{\Upsilon} = \Upsilon + \varepsilon(\Upsilon)$, where the error term $\varepsilon(\Upsilon)$ is the sum of two components

i) the bias term $b(\theta) = \dfrac{K}{2 \sum_k h_k}$

ii) a mean zero random component with variance $\dfrac{\Upsilon}{\sum_k h_k}$

Asymptotically as $\sum_k h_k$ grows large and $\Sigma\, h_k/K$ grows large, the error $\varepsilon(\Upsilon)$ becomes small.
4. Replacing Υ by $\tilde{\Upsilon}$ in the equation for the MCSE criterion yields, because of continuity, a solution that is asymptotically equivalent to the MLE for $\theta = (\beta, \mu, \sigma_n)$.

REFERENCES

Bickel, P. J. and K. A. Doksum, Mathematical Statistics, Holden-Day, Inc., Oakland, California, 1977.

Ciervo, A. P., S. K. Sanemitsu, D. E. Snead, and R. W. Suey, User's Manual for SNAP/D: Seismic Network Assessment Program for Detection, Pacific-Sierra Research Corporation, Report 1027B, 1985.

Cox, D. R. and D. V. Hinkley, Theoretical Statistics, Chapman and Hall, London, 1974.

Hutchenson, K. D., Maximum Likelihood Magnitude Calculation for Small Seismic Events in the Soviet Union, DCS-1120-14, ENESCO, Inc., 1983.

Kelly, E. J. and R. T. Lacoss, Estimation of Seismicity and Network Detection Capability, MIT, Lincoln Laboratory, TN1969-41, September 1969.

Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill, 2nd Ed., 1984, p. 185.

Parzen, Emanuel, Stochastic Processes, Holden-Day, Inc., San Francisco, California, 1962.

Rao, C. R., "Theory of the Method of Estimation by Minimum Chi-Square," Bull. Int. Stat. Inst., Vol. 35, No. 2, 1957, p. 25-32.

Richter, C. F., Elementary Seismology, W. H. Freeman and Company, Inc., San Francisco, California, 1958.

Ringdal, F., "Maximum Likelihood Estimation of Seismic Event Magnitude from Network Data," Bull. Seism. Soc. Am., Vol. 66, 1976 pp. 789-802.

Rivers, D. W., Input Data for SNAP/D, Teledyne Geotech, Alexandria, Virginia, Technical Memorandum, August 1984.

Veith, K. F. and G. E. Clawson, "Magnitude from Short-Period P-wave Data," Bull. Seism. Soc. Am., Vol. 62, 1972, pp. 435-452.

von Seggern, D. and R. Blandford, "Seismic Threshold Determination," Bull. Seism. Soc. Am., Vol. 66, June 1976, pp. 753-788.

Zavadil, R. J., D. W. Rivers, R. R. Blandford, and M. J. Shore, Estimates of Seismic Activity in the Soviet Union Based on Maximum Likelihood Estimates of Magnitude, AFTAC-ED-83-11, Patrick AFB, Florida, 1983.